Chapter 4

Computational Identification of Potential DAPK1 Inhibitors for Targeting Neuroinflammation in Alzheimer's Disease

Chapter 4

Computational Identification of Potential DAPK1 Inhibitors for Targeting Neuroinflammation in Alzheimer's Disease

4.1. Introduction

Death-associated protein kinase 1 (DAPK1) is a 160 kDa calcium/calmodulin-dependent serine/threonine kinase that plays a pivotal role in various cellular processes, particularly in apoptosis and autophagy. DAPK1, comprising of 1430 amino acids, is primarily expressed in brain regions associated with memory and cognition, making it particularly relevant to neurodegenerative diseases like AD. The protein consists of several functional domains, including a catalytic domain crucial for its kinase activity, a Ca^{2+}/CaM autoregulatory domain that inhibits activity in the absence of calmodulin, and extra-catalytic regions such as ankyrin repeats that facilitate protein interactions. DAPK1's intrinsic GTPase activity, facilitated by its ROC-COR domain, further modulates its kinase activity and interactions with other proteins. Its C-terminal death domain mediates interactions with apoptosis-related proteins, emphasizing its multifaceted role in regulating cell death and survival. Dysregulation of DAPK1 is implicated in several pathological conditions, particularly in the context of AD, where its overexpression is linked to neurodegeneration.

In the context of AD, DAPK1 is significantly involved in neuroinflammation, which is a key feature of the disease's pathology. DAPK1 has been shown to activate caspase-1 in inflammasomes, facilitating the production of pro-inflammatory cytokines such as IL-1β. This inflammatory response exacerbates neuronal damage and contributes to cognitive decline. Studies indicate that inhibiting DAPK1 can mitigate Aβ-induced memory

deficits, highlighting its potential as a therapeutic target for managing neuroinflammation associated with AD. Furthermore, the heightened expression of DAPK1 in the hippocampus of AD patients correlates with increased inflammation and neuronal loss, establishing a critical link between DAPK1 activity and the neuroinflammatory processes that underlie the progression of AD.

In the present study, an integrated ligand-based and structure-based drug design approach was employed to identify inhibitors of DAPK1. Initially, the pharmacophoric features of a known small molecule inhibitor of DAPK1 were mapped to develop a pharmacophore hypothesis. The efficacy of these models was assessed using the enrichment factor (EF) and goodness of hit (GH) scores. Subsequently, the selected pharmacophore models were utilized to screen the Zinc15 compound library, resulting in the identification of potential hits. These hits underwent drug-likeness and PAINS filtering to ensure their suitability for further investigation. Following the initial screening, a docking study was conducted in three distinct steps to yield compounds with favourable binding energies and optimal ligand-target interactions. Ultimately, three promising hits were obtained, which were subjected to rigorous molecular dynamics simulations. All the three compounds demonstrated optimal stability under the simulated conditions, along with low predicted toxicity. These identified compounds hold significant promise as potential leads for DAPK1 inhibition, contributing to the ongoing efforts to develop effective therapeutics for AD.

4.2. Materials & methods

4.2.1. Pharmacophoric feature mapping and pharmacophore generation and validation

A high-resolution (resolution 1.95 Å) X-ray crystal structure of the DAPK1 kinase domain in complex with a small molecule inhibitor (PDB ID 4TXC), obtained from the RCSB PDB website (https://www.rcsb.org/), was selected for pharmacophore hypothesis generation. **Figure 4.1 (a)** illustrates the structure of the protein–ligand complex visualised in UCSF Chimera UCSF Chimera (https://www.cgl.ucsf.edu/chimera/) [204].

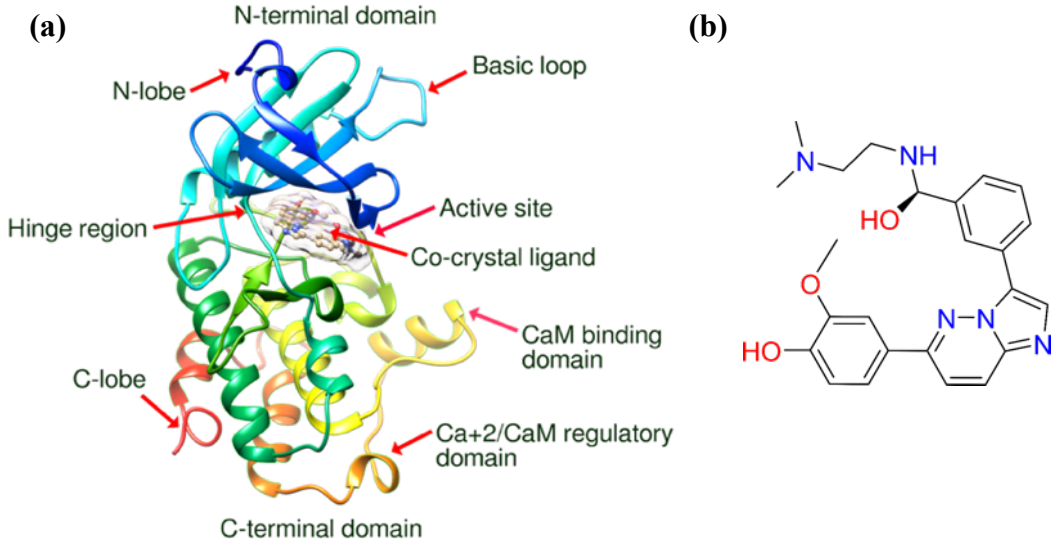


Figure 4.1 (a) DAPK1 kinase domain in complex with a small molecule inhibitor (PDB ID-4TXC), **(b)** Chemical structure of compound 38G (PDB ID-4TXC)

The co-crystallized ligand [4-(3-f3-[(R)-f[2-(dimethylamino)ethyl]amino(hydroxy) methyl]phenylgimidazo[1,2-b]pyridazin-6-yl)-2-methoxyphenol] (38G, IC_{50} 240 nm) is illustrated in **Figure 4.1 (b)**. This compound was identified as a DAPK1 inhibitor through a screening process that utilized a library of small molecules predicted to have a high affinity for binding to the conserved ATP-binding pocket of protein kinases [205]. Protein-ligand interactions and pharmacophoric features were analysed using BIOVIA

Discovery Studio Visualizer. Nine possible pharmacophore sites were identified, including hydrogen bond donor (HD), hydrogen bond acceptor (HA), aromatic group (A), and hydrophobic centre (HY). The Pharmit server (https://pharmit.csb.pitt.edu/) [206] was used to generate 15 pharmacophore hypotheses based on different combinations of the identified features.

4.2.2. Validation of ligand-based pharmacophore

To validate the pharmacophore hypotheses, EF and GH were calculated by using a decoy dataset consisting of 800 compounds on the Pharmit server. Among the dataset of compound , 41 were active compounds, while the remainder were drug-like decoys. The formulas used for these calculations are as follows:

$$EF = \frac{Ha \times D}{Ht \times A}$$

$$\mathrm{GH} = \mathrm{Ha}\,(3\mathrm{A} + \mathrm{Ht}) \left[1 - \frac{(\mathrm{Ht} - \mathrm{Ha})}{(\mathrm{D} - \mathrm{A})} \right]$$

Where 'D' represents the total number of compounds in the dataset, 'A' denotes the total number of active compounds, 'Ht' indicates the total hits, and 'Ha' is the number of active hits. Based on the calculated EF and GH values, the three most effective models were selected for pharmacophore-based virtual screening.

4.2.3. Pharmacophore-based virtual screening of a database

Pharmacophore-based virtual screening was conducted using the Pharmit server. The three pharmacophore queries generated were utilized to screen the Zinc15 database, which contained 123,399,574 conformers of 13,190,317 molecules (updated on June 2, 2019). The pharmacophoric features, including HD, HA, A, and HY, were mapped from the crystal structure of DAPK1.

The Pharmit server utilized the identified pharmacophoric features as inclusive spheres to conduct the virtual screening of the Zinc database. During the screening process, molecules were aligned with the pharmacophores to minimize the root-mean-square deviation (RMSD) between the queried features and the hit compounds. The virtual hits were identified using an RMSD cut-off value of 0.5 Å. The top ten thousand virtual hits were selected for subsequent analysis from each pharmacophore model.

4.2.4. Drug-likeliness and pan-assay interference compounds (PAINS) filtering

The KNIME Analytics Platform (version 4.2.2), an open-source software developed by [207], was employed to perform drug-likeness filtering to isolate drug-like molecules from the virtual hits. A comprehensive workflow was created using various KNIME nodes to identify the desired compounds (**Figure 4.2**). Two main filters were applied to obtain 'drug-like' compounds from the virtual hits: Lipinski's filter and the Rapid Elimination of Swill (REOS) filter (**Table 4.1**).

Table 4.1 The parameters of the Lipinski's rule of five and REOS filter

Physicochemical properties	**Acceptable limits**	
	Lipinski's rule of five	REOS
Molecular weight	Between 200 and 500	Between 200 and 500
LogP	Between -5.0 and +5.0	Between -5.0 and +5.0
H-bond donor count	Between 0 and 5	Between 0 and 5
H-bond acceptor count	Between 0 and 10	Between 0 and 10
Rotatable bond count	Between 0 and 8	Between 0 and 8
Formal charge		Between -2 and +2
Heavy atom count		Between 15 and 50

The "Lipinski Rule of Five" was utilized to filter the libraries and eliminate molecules predicted to possess poor drug-like properties, thereby conserving valuable resources in the drug development process [208]. A numerical row filter was implemented to select molecules with one or zero violations of this rule.

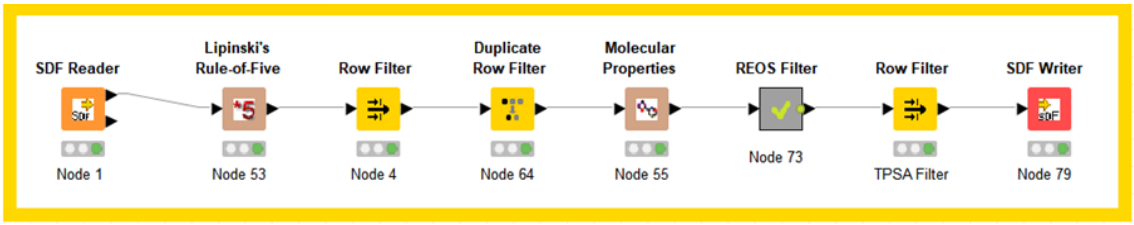


Figure 4.2 Knime workflow for drug-likeliness filtering

The REOS filter, developed at Vertex (Boston, MA), combines a rules-based approach based on physicochemical properties (including molecular weight, lipophilicity, and hydrogen-bonding characteristics) with a list of undesirable chemical groups to identify compounds [209]. This filter flags compounds containing functional groups that could lead to false positives due to reactivity, assay interference, or inadequate ADMET properties [210]. Additionally, a row filter was applied to screen for molecules with a topological polar surface area (TPSA) of less than 70 Å². The filtered molecules were then exported in SDF format using the SDF writer node. PAINS filtering was also conducted on the KNIME Analytics Platform through a separate workflow (**Figure 4.3**).

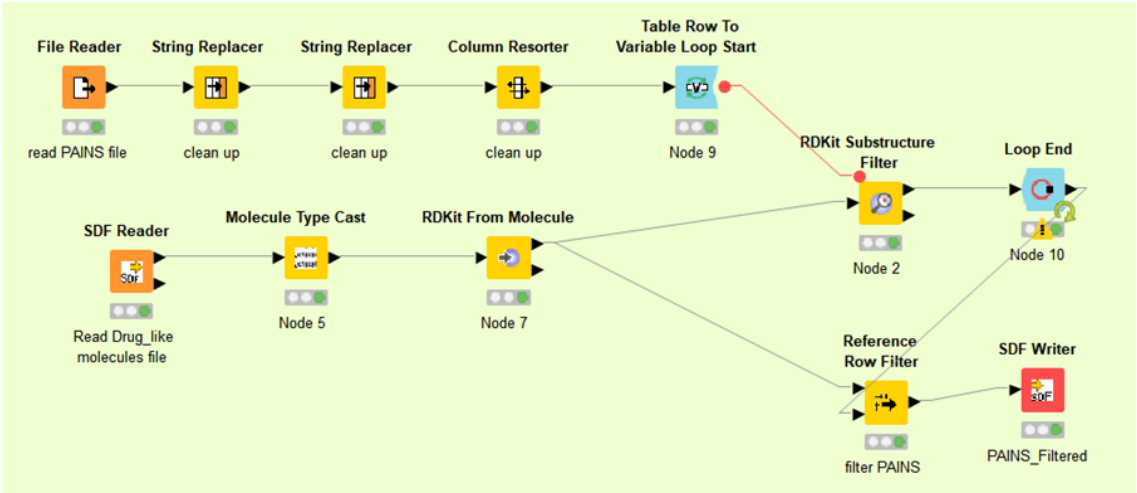


Figure 4.3 Knime workflow for PAINS filtering

4.2.5. Duplicate removal based on InChI

Duplicate ligands, represented by multiple SMILES strings, were removed by using the IUPAC Chemical Identifier (InChI) as the descriptor in the "remove duplicates by descriptor" feature of Open Babel GUI (version 3.1.1) [211].

4.2.6. Protein preparation

The X-ray crystallographic structure of the DAPK1 kinase domain (PDB ID 4TXC) was examined for any missing residues using UCSF Chimera (version 1.14) and subjected to homology modelling.

Homology modelling is a technique used to predict the three-dimensional (3D) structure of a protein based on its amino acid sequence. This method operates on the premise that proteins with evolutionary similarities often share comparable structures. Therefore, the structure of a target protein can be modelled using the known 3D structure of a related template [212]. In this study, homology modelling was employed to identify and fill the missing residues. The protein structure was analysed using the MolProbity web application [213]. A homology model was generated with SWISS-MODEL, a fully automated and web-based service specialized in protein structure homology modelling [214]. The "user template mode" was utilized to create the model using the protein sequence sourced from UniProtKB (UniProt ID P53355) and the minimized PDB structure (PDB ID 4TXC) was uploaded as the template. The quality of the homology model was assessed using global model quality estimation (GMQE) score, which integrates properties from the target-template alignment and the template structure through a multilayer perceptron. The qualitative model energy analysis (QMEAN) provides a composite score based on statistical potentials of mean force and the consistency of the model with sequence-predicted structural features. Further

modifications were made to the protein using AutoDock (version 1.5.6) to incorporate hydrogens, assign atom types, and apply Gasteiger charges. Finally, the protein was converted into PDBQT format [215].

4.2.7. Ligand preparation for high-throughput virtual screening and docking studies

High-throughput virtual screening (HTVS) was conducted using PyRx (Python Prescription, version 0.8) software. HTVS is an effective method for assessing the binding modes and affinities of a vast library of compounds to a specific receptor, aiding in the identification of potential drug candidates. The software utilizes AutoDock, Dock, and Vina as its docking tools and MGLToolsis integrated into the AutoDock functionality within PyRx [216].

Before docking, the 3D structures of the ligands were energy-minimized to optimize their configurations. The universal force field (UFF) was applied, employing the steepest descent algorithm with a total of 500 steps for energy minimization. The minimized ligands were then converted into PDBQT format suitable for HTVS. For the molecular docking studies, AutoDock (version 4.2) was utilised, and the ligands were converted into PDBQT format using AutoDockTools (version 1.5.6).

4.2.8. Grid generation and validation

The BIOVIA Discovery Studio Visualizer was employed to identify the amino acid residues involved in non-covalent interactions with DAPK1. Using these residues as a reference, two grid boxes were created around the active sites in PyRx and AutoDock.

In PyRx, the grid dimensions were configured as follows: X: 26.906, Y: 25.975, and Z: 23.152, with the centre positioned at X: 62.079, Y: 89.326, and Z: 13.371. For AutoDock, the grid was defined to encompass 56 × 50 × 46 (xyz) points, featuring a grid spacing of

0.425 Å, and the centre was set at coordinates X: 70.049, Y: 86.417, and Z: 12.073. Additionally, the grid's accuracy was validated by redocking the compound 38G, and calculating the RMSD values based on the heavy atoms between the experimentally determined co-crystallized ligand structure and the docked conformation in the BIOVIA Discovery Studio Visualizer.

4.2.9. HTVS and docking studies

PyRx was used to conduct the HTVS. The result analysis focused on filtering ligands with binding energies of -10 kcal/mol or lower, which were then subjected to additional docking studies using AutoDock (version 4.2). The docking process involved three successive stages of precision docking, implemented through the Lamarckian Genetic Algorithm (LGA) 4.2. In these stages, the number of genetic algorithm (GA) runs was set to 10, 50, and 100, respectively. A summary of the parameters used in the various stages of the docking studies is presented in **Table 4.2**.

Table 4.2 Summary of docking parameters used in AutoDock 4.2.

Stage of docking study	GA population size	Genetic algorithm run	Maximum number of energy evaluation
Stage 1	150	10	25,00,000
Stage 2	150	50	25,00,000
Stage 3	150	100	2,50,00,000

4.2.10. *In silico* ADMET prediction

The pharmaceutical industry has embraced the mantra of "fail early, fail cheap" since the 1990s, as potency alone is insufficient for successful drug development. This approach has led to the incorporation of *in silico* and *in vitro* ADMET filters during the early phases of drug discovery and development. Consequently, there has been a significant reduction in project termination rates due to ADME-related issues [217].

The compounds identified from the third stage of docking were evaluated for the ADME properties using the SwissADME server (https://www.swissadme.ch/) [218], while toxicity predictions were conducted through the PreADMET server (https://preadmet.bmdrc.kr/toxicity/).

4.2.11. Molecular dynamics simulation study

Compounds exhibiting favourable ADMET properties were subjected to molecular dynamics (MD) simulations, which is often termed the "computational microscope" [219]. It utilizes classical mechanics principles to calculate the motion of particle-based systems, which correspond to atoms or groups of atoms in biomolecules. These simulations can reveal the dynamic interactions between drugs and their protein receptors, as well as capture subtle structural changes they induce. The integration of MD and docking methodologies enhances the effectiveness of drug screening efforts [220].

MD simulations were conducted for the DAPK1 protein complexed with the selected compounds identified in the docking studies, as well as for the protein structure alone, using the Amber 20 software. The *tleap* module of Amber 20 was utilized to generate the topologies and coordinates for both the protein and the protein-ligand complexes [221]. The refined system was solvated within a cubic box filled with TIP3P water molecules, maintaining a cut-off distance of 12 Å to create an aqueous environment. Counter ions were incorporated into the system using the 'Monte Carlo' method to neutralize any residual charges. Long-range electrostatic interactions were managed using the 'Particle Mesh Ewald' (PME) method. Subsequently, the systems underwent a series of steps, including energy minimization, heating, density equilibration, and equilibration under periodic boundary conditions. The final 100 ns of MD production simulations were conducted at a temperature of 310.15 K within an NPT ensemble under periodic boundary

conditions [222]. The quality of the MD simulations was assessed using metrics such as RMSD, root mean square fluctuation (RMSF), hydrogen bonds, radius of gyration (Rg), and radial distribution functions. Post-MD analysis was carried out using the *cpptraj* module of Amber 20. Binding free energy calculations for the protein-ligand complexes were performed using the molecular mechanics/Poisson-Boltzmann surface area (MM-PBSA) method, utilizing the MMPBSA.py script of Amber 20. These calculations were conducted over the last 50 ns of the simulation at 10 ns intervals [223].

4.3. Result and discussion

4.3.1. Pharmacophore modelling and validation

A total of 15 pharmacophore models were generated based on the interactions between 38G and DAPK1 protein structure. The number of hits retrieved from screening the Zinc database is presented in **Table 4.3**. Among the 15 pharmacophore models, models 7, 10, and 15 (**Figure 4.4**) exhibited the highest efficacy values (EF and GH) as shown in **Table 4.3**.

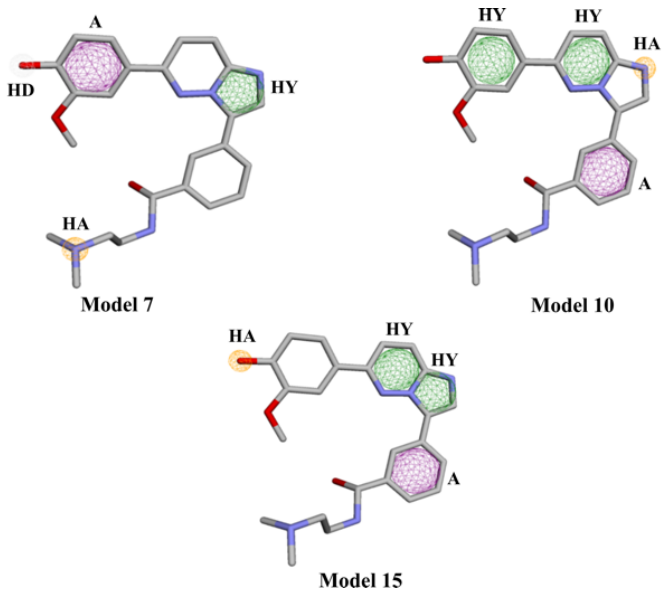


Figure 4.4 Selected pharmacophore models with pharmacophoric features (HA: H-bond acceptor, HY: hydrophobic, A: Aromatic, HD: H-bond donor)

Table 4.3 The features as well as EF and GH scores of the pharmacophore models of DAPK1 inhibitor

Pharmacophore model	Features	Hits	EF	GH
1	1HD, 1HA, 1A, 2HY	4524	*	*
2	1HD, 2A, 2HY	3657	0.798	0.141
3	2HA, 2A, 1HY	302	0	0
4	2HA, 1A, 2HY	990	*	*
5	2HA, 1A, 2HY	2667	0	0
6	1HA, 1A, 2HY	23489	0.966	0.128
7	**1HD, 1HA, 1A, 1HY**	**20009**	**1.055**	**0.463**
8	1HA, 1A, 2HY	24558	0.495	0.186
9	1HD, 1HA, 1A, 1HY	7228	0.642	0.086
10	**1HA, 1A, 2HY**	**23856**	**1.264**	**0.377**
11	1HD, 1HA, 3HY	1407	0	0
12	1HA, 1A, 2HY	22891	0.847	0.045
13	1HD, 1A, 3HY	1922	0.881	0.0273
14	1HD, 2HA, 1A, 1HY	946	*	*
15	**1HA, 1A, 2HY**	**27364**	**3.484**	**0.234**

* No molecules were filtered.

4.3.2. Drug-likeliness filter

A workflow was developed using the KNIME analytics platform to filter the selected compounds for drug-likeness. This workflow was utilized to calculate various molecular properties of the compounds, such as molecular weight, number of rotatable bonds, xLogP, TPSA, and the number of hydrogen bond acceptors (HBAs) and hydrogen bond donors (HBDs). The diverse distribution of these compounds is represented in histograms (**Figure S1**).

The filtering process incorporated 'Lipinski's Rule of Five' and the 'REOS filter' to identify compounds with optimal physicochemical properties from the results of virtual screening. Out of the 30000 compounds screened virtually, 1129 were expected to be active in the CNS after passing these filters.

4.3.3. Duplicate removal based on InChI

A chemical identifier serves as a textual label that denotes a specific chemical substance, ensuring that distinct substances are represented by unique labels [224]. The simplified molecular input line entry system (SMILES) is a notation system that encodes the molecular structures of various chemicals using concise strings of American standard code for information interchange (ASCII) characters [225]. However, multiple SMILES strings can often correspond to the same molecular structure, as the representation can begin with any atom in the molecule. In contrast, the international chemical identifier (InChI) is designed to provide a standardized structure representation, approved and recognized by IUPAC, and is both hierarchical and layered in its format [224].

To eliminate duplicate ligands that were represented by various SMILES strings, the Open Babel GUI was utilized, employing InChI as the descriptor. After the removal of duplicates, a total of 1099 unique compounds were obtained from the initial 1129 hits.

4.3.4. Homology modelling and validation

The structure of the co-crystallized ligand (PDB ID 4TXC) was analysed using the MolProbity web application, a web server dedicated to validating protein and nucleic acid structures through comprehensive evaluations of model quality at both global and local scales [226]. The SWISS-MODEL web-based application was utilized for homology modelling, employing ProMod3 (version 3.1.1) to derive initial structural data from the template structure [227]. ProMod3 addresses unfavourable interactions, minor structural distortions, and steric clashes by utilizing the CHARMM27 force field [228]. Built upon OpenStructure (OST) data structures for molecular structure manipulation [227], ProMod3 is primarily implemented in efficient C/C^{++} [229]. Its modular architecture incorporates Python scripting, facilitating flexible modelling workflows and rapid

prototyping [227]. The RMSD between the co-crystal protein structure and the homology-modelled structure was calculated using Maestro, yielding a value of 0.0594 (**Figure S2**).

The GMQE score, which indicates the anticipated accuracy of a constructed model, ranges from zero to one, with higher scores signifying greater reliability. The modelled protein's GMQE score was determined to be 0.04. The QMEAN Z-score provides an absolute assessment of the quality of protein models [230]. A QMEAN Z-score close to zero is considered favourable, while scores of -4.0 or lower suggest low-quality models. The QMEAN Z-score for the model in question was found to be 0.00. A summary of the MolProbity results for the protein structure before and after homology modelling is presented in **Table 4.4**.

Table 4.4 Structure validation results from MolProbity before and after Homology modelling

Score	Ideal Case	Before Homology modelling	After Homology modelling
Clash Score	Zero	4	1.10
Ramachandran Favoured	>98%	97.08%	97.45%
Ramachandran Outliers	<0.05%	0.36%	0.00%

4.3.5. HTVS and docking studies

The grids for HTVS and docking studies were prepared by analysing the interactions shown in BIOVIA Discovery Studio Visualizer, which determined the appropriate size of the grid box. Validation of the grid yielded a RMSD of 1.17 Å between the co-crystallized (PDB ID 4TXC) and the redocked compound 38G, indicating minimal deviation and suitability for further docking studies (**Figure 4.5**).

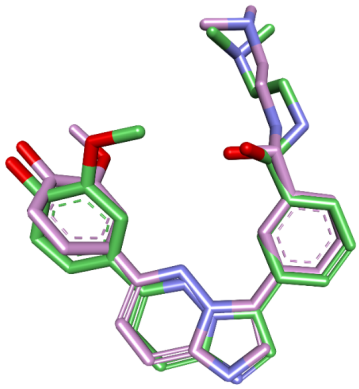


Figure 4.5 Superimposition of co-crystallised (green) and docked 38G (lavender)

The filtered ligands were then docked against the target protein. HTVS was performed using PyRx, identifying 95 compounds with binding energies of -10 kcal/mol or lower from the initial 1099 compounds. These selected ligands underwent additional molecular docking assessments using AutoDock (version 4.2) with the LGA. Binding energy was calculated using the AutoDock scoring function, and the "ligand-binding efficiency" (the ratio of binding energy to the number of heavy atoms in the molecule) was determined as an important metric. The depth of the docking studies depended on several parameters, including the number of GA runs, the maximum number of energy evaluations, and the population size [228].

From the 95 compounds resulting from HTVS, the initial stage of docking produced 30 ligands, which were further narrowed down to 10 through stage 2 docking. All 10 compounds underwent a final assessment in stage 3 docking (**Table 4.5**).

Table 4.5 Number of compounds subjected to HTVS and different stages of docking

Type of docking study	Platform used	Number of compounds subjected to docking
HTVS	PyRx 8.0	1099
Stage 1	AutoDock 4.2	95
Stage 2	AutoDock 4.2	30
Stage 3	AutoDock 4.2	10

Following this, the compounds were evaluated for *in silico* pharmacokinetics and toxicity predictions. The mean binding energy, percentage of cluster size, and ligand binding efficiency for the ligands identified in stage 3 docking studies are summarized in **Table 4.6**.

Table 4.6 Third stage docking results displaying mean binding energy, cluster size percentage, and binding efficiency of ligands

Zinc database ID	Mean binding energy (Kcal/mol)	Cluster size percentage	Ligand binding efficiency
ZINC000019799129	-8.72	37	-0.311
ZINC000004944839	-8.59	100	-0.277
ZINC000006755051	-8.24	79	-0.266
ZINC000020648330	-8.22	96	-0.274
ZINC000020622224	-8.08	42	-0.253
ZINC000067262227	-7.61	58	-0.282
ZINC000020650468	-7.57	62	-0.261
ZINC000064894179	-7.28	25	-0.214
ZINC000058405709	-7.25	83	-0.29
ZINC000009529011	-6.92	94	-0.247
38G	-7.4	76	-0.217

In the context of protein-ligand docking, mean binding energy indicates the average energy associated with the interaction between a protein and a ligand within a docking model, reflecting the combined strength and stability of binding interactions. Ligand binding efficiency assesses how effectively a ligand binds to a target protein relative to its molecular weight, highlighting its capability to achieve strong binding affinity while minimizing molecular size. Cluster size percentage refers to the fraction of generated docking poses that fall within a defined spatial grouping or cluster based on structural similarity, which assists in evaluating docking reliability and identifying potential binding modes. Additionally, the 2D and 3D interaction diagrams for the 10 ligands from stage 3

are presented in **Figure S3**, and details of the interacting amino acid residues and types of interactions are reported in **Table S1**.

4.3.6. *In silico* ADME predictions

SwissADME, a user-friendly web tool that provides free access to several efficient predictive models for physicochemical properties, pharmacokinetics, drug-likeness, and medicinal chemistry friendliness, was utilized to assess the selected compounds [218].

The analysis revealed that all selected compounds exhibit high gastrointestinal (GI) absorption, resulting in a bioavailability score of 0.55. Each of the compounds demonstrates drug-likeness, as they adhere to the criteria set forth by Lipinski's rule [231], Ghose rule [232], Veber rule [233], Egan rule [234], and Muegge rule [235]. The only exceptions were **ZINC000020622224** and **ZINC000064894179**, which each exhibited one violation of the Ghose rule. Various parameters obtained from the SwissADME web server are summarized in **Table 4.7**.

Table 4.7 The results of the ADME test with Swiss ADME

Molecule	GI absorption	BBB permeant	CYP3A4 inhibitor	CYP2D6 inhibitor	CYP2C9 inhibitor	CYP2C19 inhibitor	CYP1A2 inhibitor
ZINC000019799129	High	Yes	Yes	No	Yes	Yes	Yes
ZINC000020622224	High	Yes	Yes	Yes	Yes	Yes	Yes
ZINC000004944839	High	Yes	No	No	Yes	Yes	Yes
ZINC000020648330	High	Yes	Yes	Yes	Yes	No	Yes
ZINC000067262227	High	Yes	Yes	Yes	Yes	Yes	Yes
ZINC000020650468	High	Yes	Yes	Yes	Yes	No	No
ZINC000009529011	High	Yes	Yes	Yes	Yes	No	Yes
ZINC000006755051	High	Yes	Yes	Yes	Yes	Yes	Yes
ZINC000058405709	High	Yes	Yes	Yes	Yes	Yes	Yes
ZINC000064894179	High	Yes	Yes	Yes	Yes	Yes	No

4.3.7. *In silico* toxicity predictions

Given that all 10 compounds demonstrated high GI and BBB permeability, further toxicity assessments were conducted using the PreADMET server. The analysis revealed that all compounds were non-carcinogenic in mice. Additionally, three specific compounds, *i.e.*, **ZINC000020648330, ZINC000020650468, and ZINC000006755051**, were identified as non-mutagenic in the Ames test. These three compounds were selected for subsequent MD simulations based on these toxicity predictions and the results from stage 3 docking. **Table 4.8** summarizes the toxicity predictions for the selected compounds.

Table 4.8 Predicted *in silico* toxicities of selected compounds

ZINC ID	ZINC000020648330	ZINC000020650468	ZINC000006755051
algae_at	0.0202727	0.0151732	0.0181989
Ames_test	non-mutagen	non-mutagen	non-mutagen
Carcino_Mouse	negative	negative	negative
Carcino_Rat	positive	positive	positive
daphnia_at	0.0529693	0.0460777	0.0481491
hERG_inhibition	medium_risk	medium_risk	medium_risk
medaka_at	0.0054733	0.00415951	0.00463367
minnow_at	0.0112438	0.00682973	0.00999947
TA100_10RLI	negative	negative	negative
TA100_NA	negative	negative	negative
TA1535_10RLI	negative	negative	negative
TA1535_NA	negative	negative	negative

4.3.8. Molecular dynamics simulation study

Proteins function as dynamic entities, engaging in interactions with one another or with small molecules to carry out their biological roles or transmit signals. The interactions typically occur at specific regions on the protein surface known as hot spots, which often cluster to form binding sites that accommodate interaction partners. While not every binding site is suitable for drug development, promising ones may be concealed within

crystal structures and can be revealed by examining their dynamic properties. The flexibility of the target plays a crucial role in the formation and identification of the binding site. The flexibility of both the receptor and ligand is essential for accurately predicting ligand binding, along with associated thermodynamic and kinetic properties. Minor alterations in protein conformation, particularly the movements of side chain residues, can significantly impact the complementarity of the ligand and the protein binding site. MD simulation can assist in validating and refining docking solutions, addressing the limitations inherent in docking methods. Suboptimal docking poses produce unstable MD trajectories, potentially resulting in the ejection of ligands from their binding sites, while optimal poses maintain stability, reflected by low RMSD over time.

MD simulations can be conducted in the presence of solvents, ions, and under various physiological conditions to examine how the protein binding site interacts with water molecules and ions. Analysing MD trajectories yields insights into the binding site's persistence, varying dimensions, and the duration of hydrogen bonds. This approach effectively explores and reveals the dynamic behaviour of a target protein, facilitating the identification and analysis of specific binding sites [236].

4.3.8.1. RMSD analysis

RMSD is a measure of the average alterations in displacement of selected atoms for a specific frame with respect to a reference frame. For the Cα atoms in a protein-ligand complex, RMSD indicates how much the protein structure deviates from its initial conformation during an MD simulation. It is computed for all frames in the trajectory using the formula:

$$RMSD_X = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(r_i'(t_X)) - r_i(t_{ref})\right)^2}$$

where 'N' is the number of atoms selected, 't_{ref}' is the reference time and 'r'' is the position of the selected atoms in frame x after superimposing on the reference frame, where frame 'x' is recorded at time 't_x'. This process repeats for every frame in the simulation trajectory.

Figure 4.6 (a) presents the RMSD plot for both the native protein and the protein-ligand complexes over a 100 ns MD simulation. The native protein demonstrated consistent stability throughout the simulation, with an average RMSD of 0.83 Å (yellow). Similarly, the three protein-ligand complexes exhibited stable protein backbone conformations during the run. The average RMSD values for the complexes with **ZINC000020650468**, **ZINC000006755051**, and **ZINC000020648330** were 0.82 Å, 0.79 Å, and 0.78 Å, respectively. The ligand-bound proteins showed slightly reduced RMSD values compared to the apo form, with the **ZINC000020648330** complex showing the smallest deviation. Deviations in the range of 1–3 Å are typical for small, globular proteins, with higher values suggesting significant conformational changes. The stable RMSD values indicate that the α-carbons of the protein backbone maintained relative stability across all simulations. During the MD simulation, ligands may shift from their initial docking positions, penetrate deeper into the receptor pocket, and interact with water molecules in the binding site. If a ligand's RMSD is substantially larger than that of the protein, it could imply that the ligand has drifted away from its original binding site. However, the mean RMSD values for **ZINC000020650468**, **ZINC000006755051**, and **ZINC000020648330** were 0.71 Å, 0.78 Å, and 0.83 Å, respectively, suggesting that all

ligands remained stably bound within the protein's binding pocket and did not deviate from their initial binding sites (**Figure 4.6 (b)**).

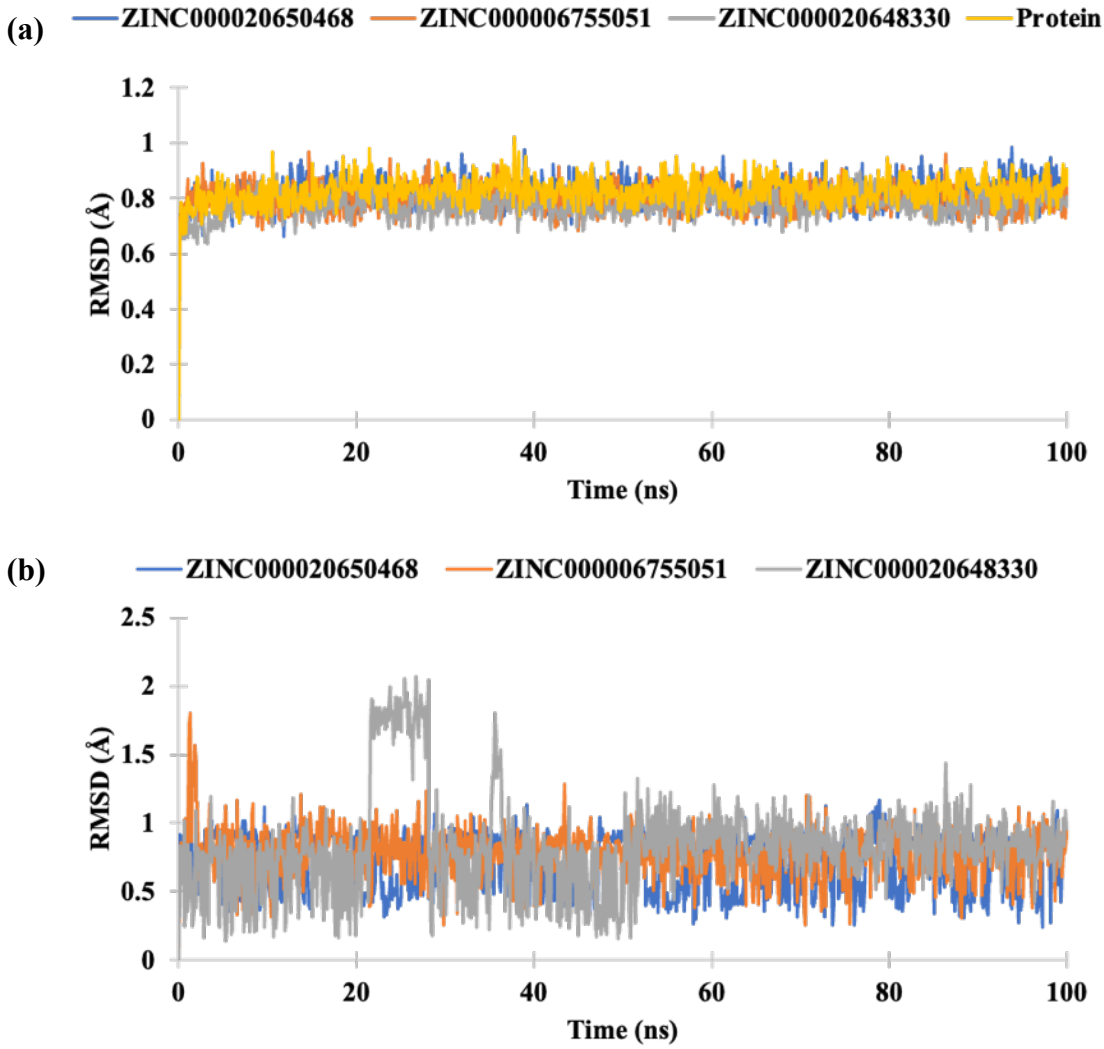


Figure 4.6 (a) RMSD of the protein backbone atoms of the protein-ligand complexes **ZINC000020650468** (blue), **ZINC000006755051** (orange), **ZINC000020648330** (grey) and the native protein (yellow). **(b)** RMSD of heavy atoms of the ligands **ZINC000020650468** (blue), **ZINC000006755051** (orange), and **ZINC000020648330** (grey).

4.3.8.2. RMSF analysis

RMSF characterizes the local changes along the protein chain. The peaks in the RMSF plot represent the residues that display the greatest fluctuations during the simulation. Typically, the tails of the protein (N- and C-terminal) fluctuate the most, while α-helices and β-strands usually fluctuate less than loop regions due to their greater rigidity. The RMSF plot indicates the stability of the secondary conformations during the MD simulation. The RMSF for the *i*th residue is calculated using the following formula:

$$RMSF_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T} < \left(r_i'(t)) - r_i(t_{ref})\right)^2 >}$$

where 'T' is the trajectory time of the RMSF calculation, 't_{ref}' is the reference time, 'r_i' designates the position of residue i, 'r'' denotes the position of atoms in residue i after superposition on the reference, and the angle brackets imply that the average of the square distance is taken over the selection of atoms in the residue.

The RMSF plot (**Figure 4.7**) illustrates stable secondary structures throughout the simulation. The average RMSF values were 0.55 Å for the native protein, and 0.58 Å, 0.59 Å, and 0.56 Å for the protein complexes with **ZINC000020650468**, **ZINC000006755051**, and **ZINC000020648330**, respectively. Notable structural fluctuations occurred at the N-terminal and C-terminal regions. Residues in the active site (specifically GLN23, PHE24, LYS42, GLU94, VAL96, GLU100, GLU146, and ASP161) exhibited good stability with lower RMSF values compared to the central loop region. High fluctuations were observed between residues 169–173, which is typical due to the presence of a beta loop. The hinge region (residues 94–96) remained stable, whereas the basic loop (residues 45–56) showed increased flexibility in both the native protein and the ligand-bound complexes. The absence of significant conformational

changes during the MD simulation suggests the homology model accurately predicts the protein's folding [228].

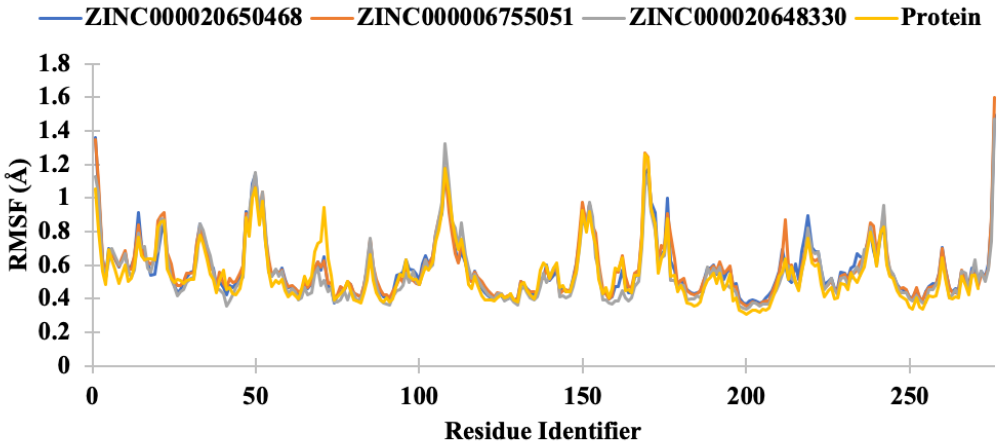


Figure 4.7 Residue-wise RMSF deviation of the protein-ligand complexes **ZINC000020650468** (blue), **ZINC000006755051** (orange), **ZINC000020648330** (grey) and the native protein (yellow).

4.3.8.3. Radius of gyration analysis

Rg refers to the distribution of a protein's atoms about its axis and measures the protein's stability. In a protein-ligand complex, calculating Rg helps researchers understand the overall compactness of the complex and how ligand binding affects the protein's structural properties. It provides insights into conformational changes or structural stability induced by ligand binding. A lower Rg value suggests that a ligand has minimally interfered with the protein's folding process, indicating a more compact structure. In contrast, a higher Rg value indicates a more loosely packed protein structure. Thus, Rg is a key parameter for analysing the conformational and folding characteristics of protein-ligand complexes. The average Rg values observed were 19.20 ± 0.054 Å for the native protein, 19.55 ± 0.048 Å for the protein–**ZINC000020650468** complex, 19.53 ± 0.051 Å for the protein–**ZINC000006755051** complex, and 19.42 ± 0.046 Å for the protein–**ZINC000020648330** complex (**Figure 4.8**).

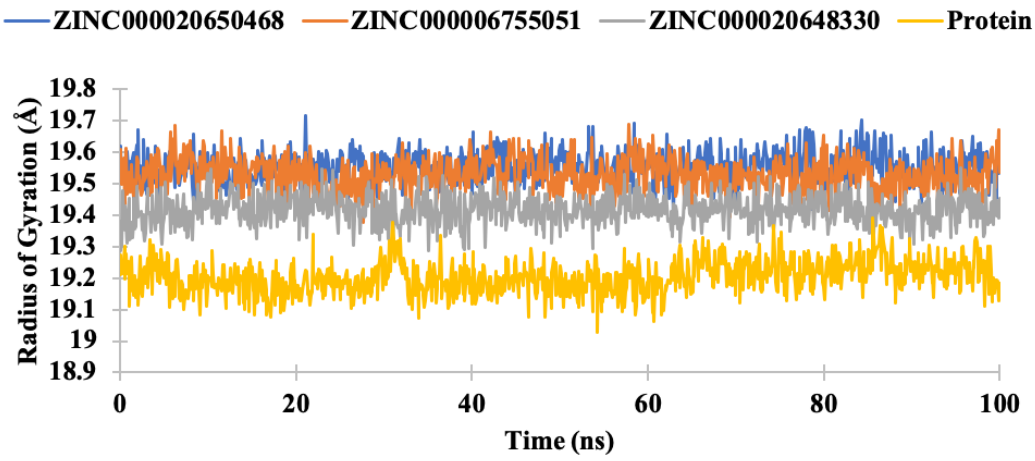


Figure 4.8 Radius of gyration of protein backbone of protein-**ZINC000020650468** complex (blue), protein-**ZINC000006755051** complex (orange), protein-**ZINC000020648330** complex (grey), and the native protein (yellow).

4.3.8.4. Hydrogen bond analysis

H-bonds are a particular kind of electrostatic interaction occurring between a proton of an electronegative atom (such as N or O) and a lone pair electron of an electronegative atom (such as N, O, or F). It is the most frequent and effective molecular interaction for stabilizing a protein-ligand complex. Following an MD simulation, an H-bond analysis is frequently performed to precisely count the number of H-bonds facilitating the protein–ligand interaction. **Figure 4.9** illustrates the number of hydrogen bonds formed throughout the simulation.

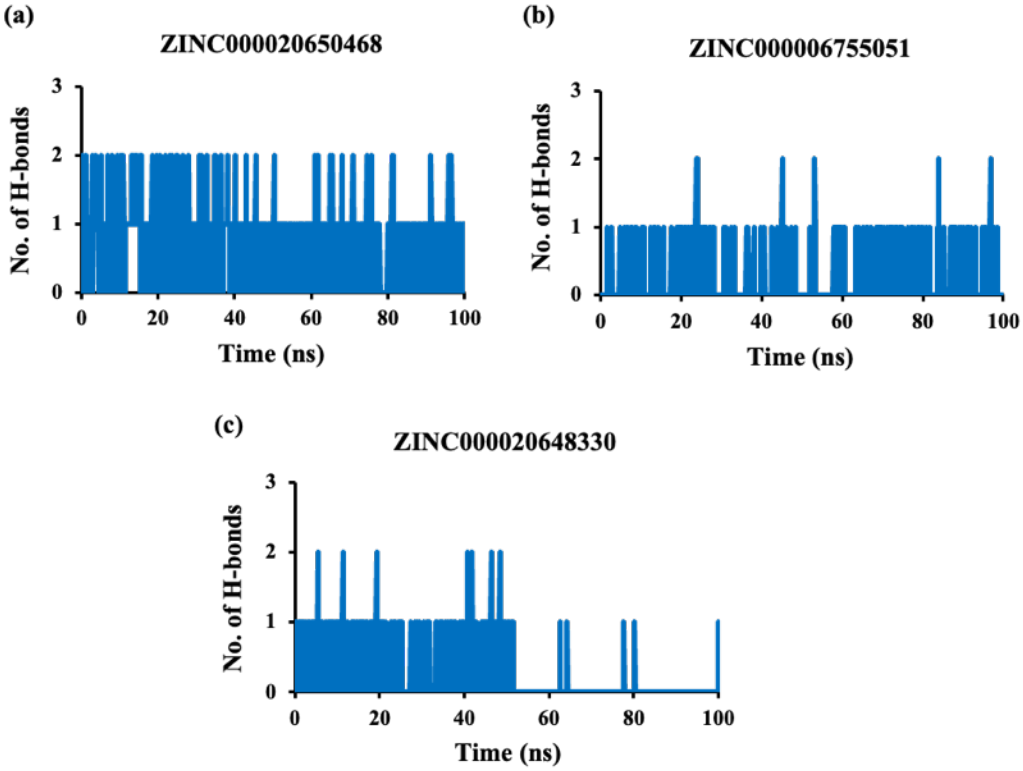


Figure 4.9 Number of intermolecular H-Bonds between ligand and protein during the entire simulation run **(a) ZINC000020650468 (b) ZINC000006755051 (c) ZINC000020648330**.

4.3.8.5. Radial distribution function analysis

The radial distribution function (RDF) is a pair correlation function that measures the density distribution of interacting radii, crucial for stabilizing the ligand at the docking site [237]. To evaluate the significance of active site residues in ligand binding, RDF plots were generated for key hydrogen-bonding residues, namely LYS42, VAL96, and ASP161.

For LYS42, the maximum density distribution points with the oxygen atoms of **ZINC000020650468**, **ZINC000006755051**, and **ZINC000020648330** were observed at 2.22 Å with g(r) = 0.20, at 2.28 Å with g(r) = 0.14, and at 2.55 Å with g(r) = 0.10, respectively (**Figure 4.10 (a)**). In the case of VAL96, the highest density was recorded for **ZINC000020650468** at 1.95 Å with g(r) = 0.124, for **ZINC000020648330** at 1.883

Å with g(r) = 0.144, and for **ZINC000006755051**, two peaks were noted at 1.683 Å and 1.250 Å, both with g(r) = 0.171 (**Figure 4.10 (b)**). For ASP161 interactions, the highest peak for **ZINC000020650468** occurred at 1.617 Å with g(r) = 0.142, and for **ZINC000006755051** at 1.717 Å with g(r) = 0.077. Two peaks were noted for **ZINC000020648330** at 1.350 Å with g(r) = 0.080 and at 1.383 Å with g(r) = 0.079 (**Figure 4.10 (c)**).

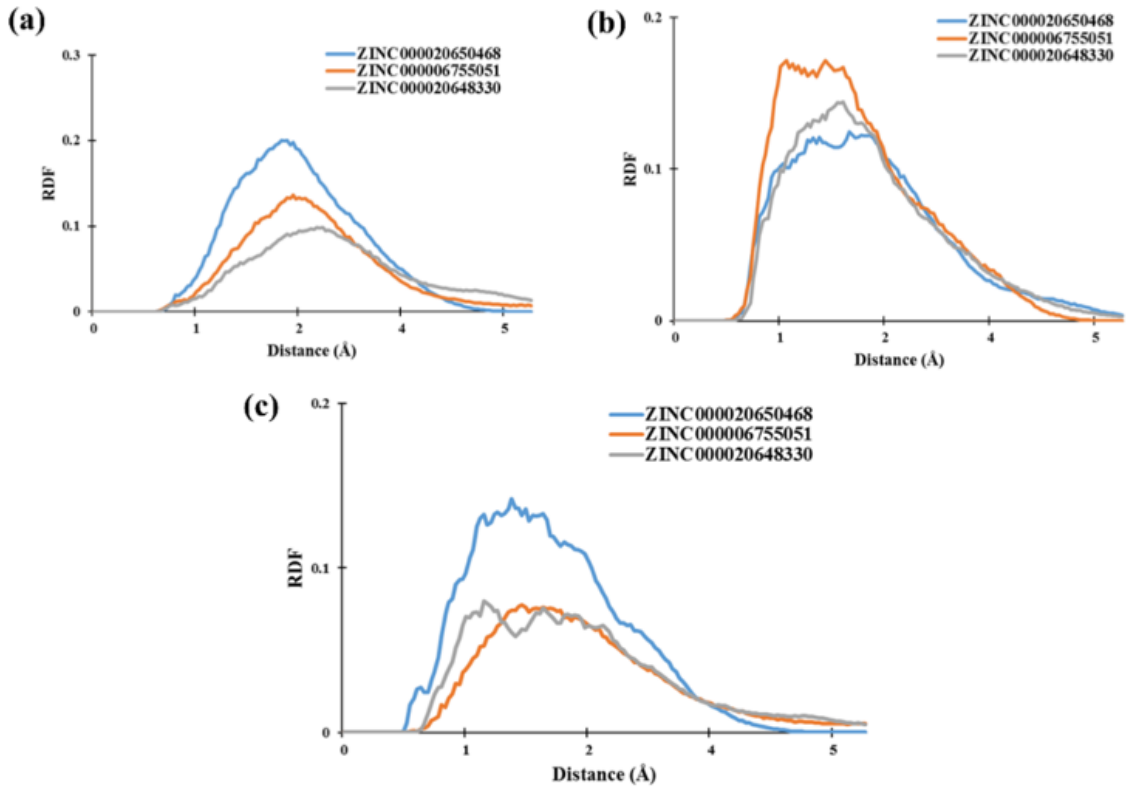


Figure 4.10 Radial Distribution Function graphs of the ligands and important residues **(a)** LYS42, **(b)** VAL96, **(c)** ASP161.

4.3.8.6. MM-PBSA assay

The binding free energy of the docked ligand-protein complexes was computed using the MM-PBSA method. **Table 4.9** presents the binding free energies of the protein-ligand complexes, along with the individual energy components: Van der Waals energy (ΔE_{vdw}), electrostatic energy (ΔE_{eel}), polar solvation free energy (ΔG_{PB}), nonpolar solvation free

energy ($\Delta G_{Nonpolar}$), and the total binding free energy ($\Delta G_{MM\text{-}PBSA}$). The ΔE_{vdw} component was identified as a key contributor to stabilizing the ligands within the protein's binding site. All three complexes demonstrated significant binding free energies, suggesting that the selected ligands form strong and favourable interactions with the protein DAPK1.

Table 4.9 Energy contributions of protein-ligand complexes in MM-PBSA assay

Compound code	ΔE_{vdw} (kcal/mol)	ΔE_{ele} (kcal/mol)	ΔG_{PB} (kcal/mol)	$\Delta G_{Nonpolar}$ (kcal/mol)	ΔG_{MMPBSA} (kcal/mol)
ZINC000020650468	-57.18±2.23	-5.14±2.09	22.82±2.74	-3.91±0.08	-43.41±2.32
ZINC000006755051	-58.38±3.72	1.16±3.27	15.94±0.98	-3.90±0.06	-45.18±4.30
ZINC000020648330	-55.29±1.83	-11.58±3.26	25.41±4.74	-3.83±0.18	-45.28± 2.12

Data expressed in mean ± S.D. (n = 6).

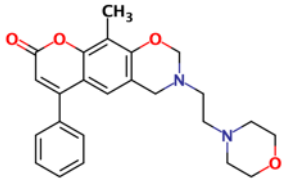
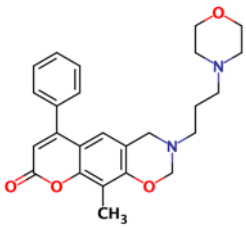
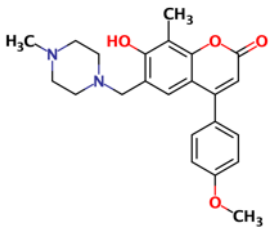
4.4. Conclusion

Computer-aided methods play a central role in modern drug discovery, from identifying starting points to guiding chemical modifications. Drug design tools enable the development of potent inhibitors for specific targets. One of the critical proteins in the pathogenesis of neurodegenerative diseases, including AD, is DAPK1, which plays a pivotal role in neuronal cell death through mechanisms like apoptosis and autophagy. Given that DAPK1 activity is significantly elevated in AD brains, inhibitors targeting DAPK1 present a promising therapeutic avenue for managing AD.

In this study, a computational molecular modelling approach was employed to identify novel DAPK1 inhibitors using a ligand-based drug design strategy. The X-ray crystal structure of the DAPK1 kinase domain bound to a small-molecule inhibitor (PDB ID: 4TXC) was used to develop three ligand-based pharmacophore models. These models were then utilized for virtual screening of the ZINC 15 database. Initially, 30000 hits were obtained, which were subsequently filtered based on drug-likeness criteria, PAINS, and

InChI-based descriptor duplicates, reducing the pool to 1099 compounds. These hits were further ranked through a combination of protein-ligand interaction studies, molecular docking, *in silico* ADMET analysis, and MD simulations. Among the identified compounds, **ZINC000020648330**, **ZINC000006755051**, and **ZINC000020650468** exhibited optimal physicochemical properties and strong protein-ligand interactions. All three compounds demonstrated favourable ADMET profiles, and MD simulations indicated that the protein-ligand complexes were stable under simulated fluid conditions. Consequently, the compounds **ZINC000020648330**, **ZINC000006755051**, and **ZINC000020650468** are proposed as potential lead compounds and are anticipated to exhibit favourable properties in subsequent experimental evaluations. The 2D structures, IUPAC names, and SMILES of the identified virtual hits are summarized in **Table 4.10**.

Table 4.10 ZINC ID, 2D structure, IUPAC name, and SMILES of identified hits

ZINC ID	**ZINC000020648330**	**ZINC000006755051**	**ZINC000020650468**
2D Structure			
IUPAC name	10-methyl-3-(2-morpholinoethyl)-6-phenyl-3,4-dihydro-2H,8H-pyrano[2',3':4,5]benzo[1,2-e][1,3]oxazin-8-one	10-methyl-3-(3-morpholinopropyl)-6-phenyl-3,4-dihydro-2H,8H-chromeno[6,7-e][1,3]oxazin-8-one	7-hydroxy-4-(4-methoxyphenyl)-8-methyl-6-((4-methylpiperazin-1-yl)methyl)-2H-chromen-2-one
SMILES	CC1=C2OC(=O)C=C(C3=CC=CC=C3)C2=CC2=C1OCN(CCN1CCOCC1)C2	CC1=C2OC(=O)C=C(C3=CC=CC=C3)C2=CC2=C1OCN(CCCN1CCOCC1)C2	COC1=CC=C(C=C1)C1=CC(=O)OC2=C(C)C(O)=C(CN3CCN(C)CC3)C=C12